Evaluation of three reference genes of Escherichia coli for mRNA expression level normalization in view of salt and organic acid stress exposure in food

Peng, Silvio ; Stephan, Roger ; Hummerjohann, Jörg ; Tasara, Taurai

Abstract: Escherichia coli can adapt to various stress conditions encountered in food through induction of stress response genes encoding proteins that counteract the respective stresses. To understand the impact and the induction of these genes under food-associated stresses, changes in the levels of their mRNA expression in response to such stresses can be analysed. Relative quantification of mRNA levels by reverse transcription quantitative real-time polymerase chain reaction (RT-qPCR) requires normalization to reference genes with stable expression under the experimental conditions being investigated. We examined the validity of three housekeeping genes (cysG, hcaT and rssA) among E. coli strains exposed to salt and organic acid stress. The rssA gene was shown to be the most stably expressed gene under such stress adaptation experimental models. The cysG gene was the least stable, whereas the hcaT gene showed similar interstrain variability as rssA but lower expression stability in the different stress adaptation models.

DOI: https://doi.org/10.1111/1574-6968.12447

RESEARCH LETTER

Evaluation of three reference genes of *Escherichia coli* for mRNA expression level normalization in view of salt and organic acid stress exposure in food

Silvio Peng[1,2], Roger Stephan[1], Jörg Hummerjohann[2] & Taurai Tasara[1]

[1]Institute for Food Safety and Hygiene, Vetsuisse Faculty University of Zurich, Zurich, Switzerland; and [2]Research station Agroscope Liebefeld-Posieux ALP, Bern, Switzerland

Correspondence: Taurai Tasara, Institute for Food Safety and Hygiene, Vetsuisse Faculty University of Zurich, Winterthurerstrasse 272, CH-8057 Zurich, Switzerland. Tel.: + 41 44 6358651; fax: + 41 44 6358908; e-mail: tasarat@fsafety.uzh.ch

Received 11 December 2013; revised 17 April 2014; accepted 17 April 2014. Final version published online 14 May 2014.

DOI: 10.1111/1574-6968.12447

Editor: Robert Gunsalus

Keywords
gene expression; organic acid stress; reference gene; RT-qPCR; salt stress; stress response.

Abstract

Escherichia coli can adapt to various stress conditions encountered in food through induction of stress response genes encoding proteins that counteract the respective stresses. To understand the impact and the induction of these genes under food-associated stresses, changes in the levels of their mRNA expression in response to such stresses can be analysed. Relative quantification of mRNA levels by reverse transcription quantitative real-time polymerase chain reaction (RT-qPCR) requires normalization to reference genes with stable expression under the experimental conditions being investigated. We examined the validity of three housekeeping genes (*cysG*, *hcaT* and *rssA*) among *E. coli* strains exposed to salt and organic acid stress. The *rssA* gene was shown to be the most stably expressed gene under such stress adaptation experimental models. The *cysG* gene was the least stable, whereas the *hcaT* gene showed similar interstrain variability as *rssA* but lower expression stability in the different stress adaptation models.

Introduction

Escherichia coli adapts to various stress conditions encountered in food, including low water activity and the presence of organic acids (Peng *et al.*, 2011). When exposed to such stresses this bacterium responds by inducing the expression of stress response genes encoding protective mechanisms, which counteract the stress and allow survival in harsh environments. To improve our understanding of *E. coli* stress responses, the impact of food-related stress conditions on the expression of stress response genes in this organism need to be investigated. Gene expression analysis based on relative quantification of mRNA levels by RT-qPCR requires normalization to cater for variation related to RNA purification, complementary DNA (cDNA) synthesis, and qPCR (Bustin, 2000). For this purpose, internal reference genes, such as housekeeping (HKG) genes, are selected and used (Bustin, 2000; Vandesompele *et al.*, 2002). Since such genes can also be regulated and might vary between different experimental conditions (Thellin *et al.*, 1999; Vandesompele *et al.*, 2002), it is important that potential reference genes are validated under specific stress conditions applied in mRNA expression studies. The aim of the present study was to validate three *E. coli* HKGs previously applied as reference genes under other experimental conditions for their suitability as internal reference genes for mRNA expression normalization under salt and organic acid stress conditions.

Materials and methods

Housekeeping genes, bacterial strains and growth

For this study three HKGs, *cysG* (encoding uroporphyrin III C-methyltransferase), *hcaT* (encoding 3-phenylpropionic transporter) and *rssA* (encoding 16S ribosomal RNA) that had previously been evaluated and used as internal reference genes under other experimental conditions were selected (Zhou *et al.*, 2011). The gene expression experiments were performed using five *E. coli* strains

which previously had been isolated from raw milk cheese: three Shiga toxin-producing (K331-4, K356, N09-1208) and two generic *E. coli* strains (FAM21843, K303). These strains were characterized *in vitro* and recently used to study *E. coli* survival during raw milk cheese production and ripening (Peng *et al.*, 2012, 2013). Primary cultures were prepared from each strain starting from single colonies by inoculating 5 mL of Luria–Bertani (LB) broth (Difco™ LB Broth, Miller, Becton Dickinson AG, Allschwil, Switzerland) and growing for 8 h at 37 °C and 200 r.p.m. Secondary cultures were prepared from the primary cultures by diluting (1 : 100) into 5 mL of LB and growing 16–18 h at 37 °C and 200 r.p.m. To prepare stationary growth phase stage tertiary cultures subsequently used in stress exposure experiments, the secondary cultures were diluted (1 : 100) into 25 mL of LB and grown at 37 °C and 200 r.p.m. to an optical density ($OD_{590\ nm}$) of 1.50 ± 0.05, which corresponded to the stationary phase in all the examined *E. coli* strains as determined by growth curves.

Stress exposure, sampling and RNA purification

From the tertiary stationary phase cultures prepared as described above, 10-mL aliquots were centrifuged (5 min at 5000 ***g***). The recovered *E. coli* cells were resuspended using 10 mL of (i) LB (control), (ii) LB acidified to pH 5.2 using lactic acid (organic acid stress) or (iii) LB containing 5% (w/v) sodium chloride (salt stress) and incubated at 37 °C. Samples of 0.5 mL were collected from the stationary phase cultures prior to the centrifugation and from the resuspended cultures after incubation for 15 and 60 min at 37 °C. In all samples the transcript profiles were immediately stabilized using 1 mL of RNAprotect Bacteria Reagent (Qiagen AG, Hombrechtikon, Switzerland). Initially, three total RNA isolation protocols were compared to determine the protocol giving the best quantity and quality of isolated RNA. In two of these protocols, an enzymatic lysis step was applied using either lysozyme alone or a combination of lysozyme and Proteinase K as described in the RNAprotect Bacteria Reagent Handbook (Qiagen). In the third protocol, a chemical and mechanical lysis procedure previously described by Arguedas-Villa *et al.* (2010) was used. RNA was subsequently isolated from the bacterial cell lysates generated from these protocols following the RNeasy Plus Mini Kit protocol (Qiagen), which includes genomic DNA removal using the genomic DNA binding column and an on-column DNase I digestion. Total RNA was eluted in 50 μL RNAse-free water, and subsequently quantified and quality controlled using the Nanodrop and BioAnalyzer instruments, respectively. This comparison established that the combination of chemical and mechanical lysis gave the highest quantity and best quality of total RNA. Hence this lysis protocol in combination with the RNeasy Plus Mini Kit was adopted for RNA isolation in this study. Briefly, the bacterial cell pellets were resuspended in 0.5 mL of the lysis buffer provided in the RNeasy Plus Mini Kit. The samples were transferred into MagNA lyser tubes and mechanically disrupted twice (60 s at 6500 r.p.m.) in the MagNA Lyser instrument (Roche Molecular Diagnostics, Rotkreuz, Switzerland). Each strain was assayed in three independent biological assays.

cDNA synthesis and quantitative real-time PCR

For cDNA synthesis, 100 ng of each total RNA sample were reverse transcribed to cDNA using the QuantiTect Reverse Transcription Kit (Qiagen). cDNA synthesis was performed in duplicate from each RNA sample. A control omitting the reverse transcriptase was carried out for each RNA sample to rule out residual genomic DNA contamination. The cDNA samples were diluted 1 : 10 and subsequently used as qPCR templates. The qPCR reactions were carried out in the Light Cycler 480 instrument (Roche Molecular Diagnostics). Previously described qPCR primers (Zhou *et al.*, 2011) and the QuantiTect SYBR Green PCR Master Mix (Qiagen) were used. The amplification efficiency of each primer pair was determined using 10-fold dilution series of genomic DNA templates isolated from the five *E. coli* strains. PCR reactions were conducted in 10 μL using genomic DNA or cDNA as template. PCR thermocycling included a 15 min hot start at 95 °C, followed by 45 amplification cycles (10 s at 95 °C, 20 s at 50 °C, 20 s at 72 °C, 1 s at 78 °C with a single fluorescent measurement). The specificity of the amplification products was checked by a melting curve (65–97 °C at 2.2 °C s^{-1} and a continuous fluorescent measurement) analysis, as well as by loading on agarose gel. Reactions were run in triplicate and included negative (water) and no-reverse transcriptase controls.

Data analysis

Crossing point (CP) values were obtained by the 'second derivative maximum' method as applied by the LC480 software (Roche Molecular Diagnostics). The CP (also called threshold cycle) is the cycle number at which the fluorescence intensity generated in the qPCR reaction rises significantly above the threshold level. To determine the most stably expressed reference gene, the CP raw data output was subsequently analysed using the EXCEL-based program BESTKEEPER (Pfaffl *et al.*, 2004). This program compares expression stability of HKGs based on a multitude of pair-wise correlation analyses and determines the

BESTKEEPER index. The software compares each gene to this index, thus calculating the Pearson correlation coefficient (r) and the correlation probability (p) between the index and the contributing candidate HKG. It also calculates standard deviations (SD) of the CP values and indicates SD higher than the cut-off value of 1. Genes with the highest coefficient of correlation to the BESTKEEPER and SD $\leq$ 1.0 are considered to be most stably expressed (Pfaffl *et al.*, 2004).

Results and discussion

Optimal stress exposure conditions, RNA purification and RT-qPCR protocols were initially determined. Strain-specific PCR amplification efficiencies of each RT-qPCR primer pairs were determined using genomic DNA samples isolated from the five *E. coli* strains as templates. Agarose gel examination showed that all three primer pairs that were used to target the *cysG*, *hcaT* and *rssA* genes, respectively, gave single PCR amplicons of expected size. PCR amplification efficiencies determined for these primers ranged between 92.4% and 96.3%. Melting curve analysis of the amplification products in qRT-PCR gave single peaks (data not shown). All five *E. coli* strains displayed similar growth curves and growth rates in the tertiary LB cultures, which were subsequently used in stress exposure experiments and reference gene expression stability evaluation (data not shown). The bacterial cells used were derived from early stationary phase growth stage cultures ($OD_{590\ nm}$ = 1.5 and approximately 10^9 colony forming units (CFU) mL^{-1} in each strain). Viable cell counting in control and stress exposed samples showed that there was no growth in any strain during the experimental time period (60 min) used in this study.

Preliminary evaluation of different RNA purification protocols established that cell lysis using a chemical and mechanical lysis combination yielded RNA of best quality and quantity in comparison with enzymatic lysis protocols using lysozyme alone or in combination with Proteinase K (data not shown). Using this protocol, RNA templates with RNA integrity numbers (RIN) ranging between 8.0 and 9.5 were obtained, indicating that the RNA was of good quality (Schroeder *et al.*, 2006). This protocol was therefore adopted and applied in this study.

The expression of *rssA*, *hcaT* and *cysG* across five *E. coli* strains was initially assessed in stationary phase organisms as well as in organisms adapted to different conditions applied in experimental models that assess salt and organic acid stress exposure impacts on gene expression. The gene showing the highest transcript abundance was *rssA* (CP ranging from 8.52 to 13.07), whereas *hcaT* (CP ranging from 27.80 to 36.68) and *cysG* (CP ranging from 24.96 to 35.55) were expressed at similar but lower levels across the five strains and under all considered experimental conditions (Fig. 1). Using CP raw data, the expression stability of the three potential reference genes was next evaluated using the BESTKEEPER program.

Intrastrain expression variation comparison between independent biological replicates showed that *cysG* under salt stress and *hcaT* under acid stress conditions displayed the highest and least expression variabilities, respectively (Supporting Information, Table S1). The c*syG* expression in salt stress exposed N09-1208 strain had the highest (SD $\pm$ 2.02) CP variation and *hcaT* expression in acid stress exposed K303 strain cells showed the least (SD $\pm$ 0.08) CP variation between independent biological replicates. The salt and organic acid stress experimental models are designed to compare stress response gene expression levels between control *E. coli* exposed to LB broth and those in organisms exposed to salt and organic acid stress in LB, respectively. As such the interstrain expression stability of the three reference genes across the five *E. coli* strains was initially evaluated in organisms exposed to single experimental conditions of these stress adaptation models as well as among stationary phase organisms (Table 1). The *rssA* gene showed the lowest (0.46 < SD < 0.98) interstrain variation and high correlation (0.775 < r < 0.937) to the BESTKEEPER index under all conditions. The *hcaT* gene showed medium to high inter-

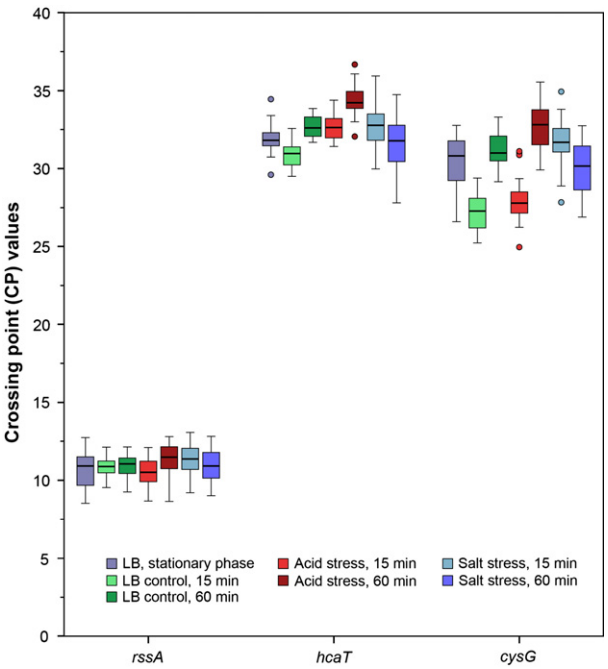

Fig. 1. CP values of three *Escherichia coli* reference genes under control, organic acid and salt stress conditions. Boxplots highlight the median, 25th and 75th percentiles and expression range across five *E. coli* strains.

Table 1. Results from BESTKEEPER descriptive statistical analysis of three reference genes across five *Escherichia coli* strains at the single control and stress conditions investigated

	LB control			Organic acid stress		Salt stress	
	stat.	15 min	60 min	15 min	60 min	15 min	60 min
rssA							
SD (± CP)	0.98	0.46	0.52	0.72	0.82	0.88	0.86
Coeff. of corr. (*r*)	0.937	0.775	0.897	0.849	0.859	0.865	0.899
P-value	0.001	0.001	0.001	0.001	0.001	0.001	0.001
hcaT							
SD (± CP)	0.64	0.66	0.57	0.71	0.75	**1.05**	**1.18**
Coeff. of corr. (*r*)	0.778	0.636	0.320	0.711	0.704	0.725	0.803
P-value	0.001	0.001	0.128	0.001	0.001	0.001	0.001
cysG							
SD (± CP)	**1.37**	0.97	0.89	**1.06**	**1.30**	**1.10**	**1.61**
Coeff. of corr. (*r*)	0.663	0.659	0.599	0.673	0.593	0.683	0.914
P-value	0.001	0.001	0.002	0.001	0.002	0.001	0.001

SD (± CP), standard deviation of the CP values (SD higher than BESTKEEPER cut-off value of 1 indicated in bold letters); coeff. of corr. (*r*) compared with BESTKEEPER index including *P*-value.

strain variation ($0.57 < SD < 1.18$), displaying an SD that was greater than the BESTKEEPER cut-off value of 1 across the five strains under salt stress. The *cysG* gene exhibited the highest interstrain variation ($0.89 < SD < 1.61$). It only displayed SDs below the BESTKEEPER cut-off value among the five strains in LB but not in the presence of salt or organic acid stress.

Next, the expression stability of these genes was examined across experimental conditions employed in the two stress adaptation models (Table 2). The *rssA* gene remained the most stable candidate reference gene ($0.62 < SD < 0.69$) and also showed high correlation to the BESTKEEPER index ($0.736 < r < 0.862$). The *hcaT* gene showed higher SDs ($0.95 < SD < 1.19$), especially in the acid stress adaptation model, where the combined SD increased to 1.19. This was because its expression decreased under acid stress as well as over exposure time [CP geometric mean at 15 min of 30.88 (LB) < 32.64 (LB acid stress) and at 60 min 32.67 (LB) < 34.31 (LB acid stress)], which indicates downregulation of transcripts of this gene due to organic acid stress and exposure time period. The SDs of *cysG* were higher than those observed for both *rssA* and *hcaT*, and higher than the BESTKEEPER cut-off value in both stress adaptation models under all conditions investigated. A similar decrease in *cysG* CP values occurred after 15 min in the LB control and the organic acid stress samples, indicating that the inoculation of stationary phase cells into fresh LB broth already led to an increased expression of the *cysG* gene.

Table 2. Results from BESTKEEPER descriptive statistical analysis of three reference genes across five *Escherichia coli* strains in organic acid and salt stress adaptation models

	Acid stress model*			Salt stress model*		
	15 min	60 min	comb.	15 min	60 min	comb.
rssA						
SD (± CP)	0.62	0.69	0.67	0.69	0.69	0.69
Coeff. of corr. (*r*)	0.736	0.840	0.737	0.766	0.862	0.790
P-value	0.001	0.001	0.001	0.001	0.001	0.001
hcaT						
SD (± CP)	0.97	0.98	**1.19**	**1.13**	0.95	**1.07**
Coeff. of corr. (*r*)	0.619	0.717	0.798	0.830	0.742	0.796
P-value	0.001	0.001	0.001	0.001	0.001	0.001
cysG						
SD (± CP)	**1.03**	**1.25**	**2.26**	**2.35**	**1.28**	**1.92**
Coeff. of corr. (*r*)	0.688	0.693	0.863	0.859	0.853	0.848
P-value	0.001	0.001	0.001	0.001	0.001	0.001

*Stress adaptation models include LB (control) and LB (stress) samples taken after the same exposure time to organic acid or salt stress as well as the combination (comb.) of the two different exposure times to a certain stress. SD (± CP), standard deviation of the CP values (SD higher than BESTKEEPER cut-off value of 1 indicated in bold letters); coeff. of corr. (*r*) compared with BESTKEEPER index including *P*-value.

In conclusion, the validity of three potential reference genes has been evaluated in LB-based salt and acid stress adaptation models. Overall, the most stably expressed candidate gene was *rssA*. Thus this gene is recommended as a reference gene for mRNA expression level normalization in LB-based salt and organic acid stress adaptation models. The interstrain variability of *hcaT* was low to medium at control and organic acid stress conditions, but higher than the BESTKEEPER cut-off value at salt stress conditions. In the stress adaptation models, the SDs of *hcaT* were close to the BESTKEEPER cut-off value or higher, indicating variations in the mRNA expression level between the conditions investigated, such as the increase in *hcaT*

CP values over organic acid stress exposure time. The expression of *cysG* gene was highly variable and seemed to be strongly regulated under the experimental conditions applied in the two LB-based stress adaptation models. As such it is recommended that this gene is not used as a reference gene in the study of stress response gene expression under either salt or acid stress conditions in LB media.

References

Arguedas-Villa C, Stephan R & Tasara T (2010) Evaluation of cold growth and related gene transcription responses associated with *Listeria monocytogenes* strains of different origins. *Food Microbiol* **27**: 653–660.

Bustin SA (2000) Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *J Mol Endocrinol* **25**: 169–193.

Peng S, Tasara T, Hummerjohann J & Stephan R (2011) An overview of molecular stress response mechanisms in *Escherichia coli* contributing to survival of Shiga toxin-producing *Escherichia coli* during raw milk cheese production. *J Food Prot* **74**: 849–864.

Peng S, Stephan R, Hummerjohann J, Blanco J & Zweifel C (2012) *In vitro* characterization of Shiga toxin-produding and generic *Escherichia coli* in respect of cheese production-relevant stresses. *Arch Lebensmittelhyg* **63**: 136–141.

Peng S, Hoffmann W, Bockelmann W, Hummerjohann J, Stephan R & Hammer P (2013) Fate of Shiga toxin-producing and generic *Escherichia coli* during production and ripening of semihard raw milk cheese. *J Dairy Sci* **96**: 815–823.

Pfaffl MW, Tichopad A, Prgomet C & Neuvians TP (2004) Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BESTKEEPER-Excel-based tool using pair-wise correlations. *Biotechnol Lett* **26**: 509–515.

Schroeder A, Mueller O, Stocker S, Salowsky R, Leiber M, Gassmann M, Lightfoot S, Menzel W, Granzow M & Ragg T (2006) The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC Mol Biol* **7**: 3.

Thellin O, Zorzi W, Lakaye B, De Borman B, Coumans B, Hennen G, Grisar T, Igout A & Heinen E (1999) Housekeeping genes as internal standards: use and limits. *J Biotechnol* **75**: 291–295.

Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A & Speleman F (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* **3**: 7. RESEARCH0034.1-0034.11.

Zhou K, Zhou L, Lim Q, Zou R, Stephanopoulos G & Too HP (2011) Novel reference genes for quantifying transcriptional responses of *Escherichia coli* to protein overexpression by quantitative PCR. *BMC Mol Biol* **12**: 18.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1. Intrastrain variations in CP values of three reference genes across five *E. coli* strains at the different control and stress conditions investigated.